Mixed Poisson distributions in exact solutions of Stochastic Auto-regulation Models

Srividya Iyer-Biswas*
James Franck Institute, University of Chicago, Chicago, IL 60637

C. Jayaprakash
Department of Physics, The Ohio State University, Columbus, OH 43210

In this paper we study the interplay between stochastic gene expression and system design using simple stochastic models of auto-activation and auto-inhibition. Using the Poisson Representation, a technique whose usefulness in the context of non-linear gene regulation models we elucidate, we are able to write down exact results for these feedback models in the steady state. We exploit this representation to analyze the parameter-spaces and demarcate where different behaviors including power-law, conventional bimodal, a novel bimodal with graded characteristics and sub-Poisson noise occur. Using our results, we reexamine how well the auto-inhibition and auto-activation models serve their conventional roles as paradigms for noise suppression and noise exploitation respectively.

PACS numbers: 87.10.Mn

I. INTRODUCTION

Stochastic fluctuations occur inevitably in the biochemical reactions that implement cell functions resulting in cell-to-cell variability of the components, in particular, in protein numbers. This noise in protein numbers can either be functionally significant or may be detrimental and thus require suppression [1]. Such fluctuations arise even in a population of cells that were initially identical due to the inherently probabilistic nature of the key biochemical processes involved in gene activation, transcription and translation coupled with the small numbers of molecules involved [2, 3]. Stochastic gene expression and specifically, such 'intrinsic' fluctuations in protein numbers, have been the focus of several experimental and theoretical studies [2–10]. An important systems biological quest is to understand if the activity of a protein is determined predominantly by just its abundance (average value) or if its cell-to-cell variability (noise strength and shape) is also an important determinant, given a gene network architecture.

Previous studies have indicated that tight control of protein numbers, when desired, is often achieved by an auto-repression motif of gene expression [11–15]. It has even been argued that the reason why this motif occurs far more frequently in nature (40% of the known transcription factors in *E. Coli* are controlled by negative auto-regulation [16]) than in studies of randomized networks is because it achieves stability against fluctuations [17]. On the other hand, noise may be exploited by cells to switch between different expression states especially in key cellular processes where 'locking' of sub-populations of cells into distinct fates needs to be achieved without changing the underlying network structure. The auto-activation motif has been implicated in such systems when population heterogeneity is desirable [18–21].

It is now well established that in many systems proteins are produced in 'bursts' with significantly varying dead times in between bursts. This important aspect of gene expression is encapsulated in a model in which the gene can stochastically switch between long-lived "off" states and "on" states leading to spurts of mRNA and protein expression. The dynamics of protein bursting depends on the interplay of time scales, the switching rate and degradation rates, in such a model. We explore the interplay between stochastic gene expression and system design by studying two simple stochastic bursting models with auto-regulation, i.e., with feedback from the protein produced to the gene switching rate. In the positive (negative) feedback model, the amount of protein produced proportionally increases the propensity of the gene to be in the "on (off)" states.

We use the Poisson representation, a technique whose usefulness in the context of solving nonlinear models of gene expression we elucidate, to derive not just the exact steady state protein distributions in these models, but to also analyze the parameter spaces and to demarcate where in them bimodal, power-law tailed, sub-Poisson and other distributions occur. Using these results we re-examine how well these two models, characterized by negative and positive feedback, serve as paradigms for noise suppression and noise exploitation respectively.

While the idea of writing down protein distributions as exact linear superpositions of Poisson distributions is relatively new [9], mixtures of Poisson distributions have long been studied in various contexts including photon statistics in quantum optics [22] and in accident proneness models in actuarial sciences [23]. Remarkably, in both the auto-activation and the auto-repression cases, we find classes of mixed Poisson distributions that, to the best of our knowledge, have not been previously considered [24]. What is more, they arise *dynamically* in these models. We also show that the Beta-Poisson mixture, which has been previously studied as a versatile *prior* distribution in accident proneness models [23], arises as

*iyerbiswas@uchicago.edu

a limiting case from these distributions.

II. THEORETICAL FRAMEWORK

Detailed expositions of the Poisson Representation can be found in [23, 25] amongst others. We have briefly discussed the Poisson Representation for linear models of gene regulation without feedback such as the gene pulsing model, in [9]. We analyze how the exact steady-state protein distributions $P(p)$ of the models with feedback discussed in this paper may be represented as a superposition of Poisson distributions, with a weighting probability density $\rho(\lambda)$ for the Poisson mean λ. In other words, we examine whether a probability density $\rho(\lambda)$ can be found such that

$$P(p) = \int_0^\infty \frac{e^{-\lambda}\,\lambda^p}{p!}\rho(\lambda)\,d\lambda \tag{1}$$

If such a $\rho(\lambda)$ is found, an immediate consequence is that the corresponding $P(p)$ is super-Poisson, i.e., has a Fano-factor (FF) > 1. $\rho(\lambda)$ is a function of a continuous variable in contrast to $P(p)$ and its convexity and monotonicity properties are easier to ascertain and these in turn determine the shapes of $P(p)$. Bimodal $P(p)$ distributions correspond to concave(upwards) $\rho(\lambda)$; Power-law tails in $P(p)$ arise when $\rho(\lambda)$ itself has a monotonically decreasing power law region, and a monotonically increasing $\rho(\lambda)$ leads to unimodal P(p) distributions with a mode around the upper edge of the λ interval. When $\rho(\lambda)$ is concave downwards with a maximum at some intermediate value of λ, then unimodal P(m) distributions with a mode around the same value obtains. We use the exact, analytical expression that we derive for $\rho(\lambda)$ to map out where in the parameter space each qualitatively distinct shape arises.

For the auto-activation model schematically shown in Figure S1, let $P_0(p,t)$ and $P_1(p,t)$ denote the probabilities that there are p proteins at time t and that the gene is in the off and on state respectively. Master Equations describing the time-evolution of these probabilities can be written down using a standard procedure [25].

Let us define $\rho_0(\lambda)$ and $\rho_1(\lambda)$ as

$$P_\alpha(p,t) \;=\; \int_0^\infty d\lambda\,\rho_\alpha(\lambda,t)\,e^{-\lambda}\,\frac{\lambda^p}{p!} \text{ for } \alpha = 0 \text{ or } 1, \tag{2}$$

Note that $\rho(\lambda) = \rho_0(\lambda)+\rho_1(\lambda)$ satisfies the normalization condition $\int d\lambda\,\rho(\lambda) = 1$. The Master equations for $\rho_\alpha(\lambda)$ are give by

$$\begin{aligned}\partial_t\rho_0(\lambda,t) &= -c_f\rho_0(\lambda,t) + c_b\rho_1(\lambda,t) + \partial_\lambda[\lambda\rho_0(\lambda,t)] \\ &\quad - a\,(\lambda\rho_0 - \partial_\lambda(\lambda\rho_0)) \qquad (3a)\\ \partial_t\rho_1(\lambda,t) &= c_f\rho_0(\lambda,t) - c_b\rho_1(\lambda,t) + \partial_\lambda[\lambda\rho_1(\lambda,t)] \\ &\quad + a\,(\lambda\rho_0 - \partial_\lambda(\lambda\rho_0)) - p_b\partial_\lambda\rho_1(\lambda,t)\,, \qquad (3b)\end{aligned}$$

with the boundary condition

$$\pm\, a\;e^{-\lambda}\,\frac{\lambda^p}{p!}\,\lambda\rho_0(\lambda)\Big|_0^{\lambda_{max}} = 0 \tag{4}$$

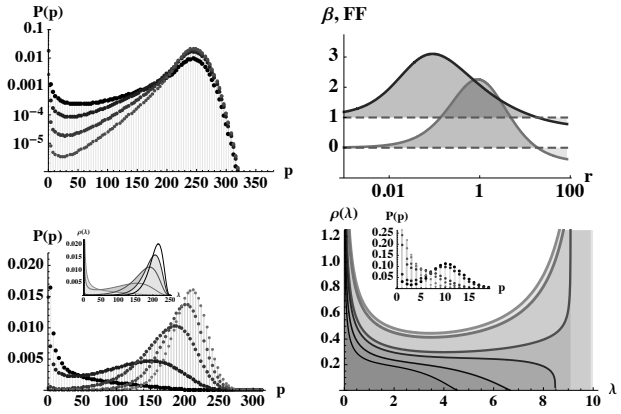

FIG. 1: *Top Left*: The response of the protein distribution to increasing activation strength in the $\phi < 1, \beta < 1$ quadrant in the auto-activation model resembles the classic 'binary response' associated with auto-activation systems. *Bottom Left*: The effect of increasing activation strength, a, on the protein distribution in the fourth quadrant, i.e., $\phi < 1, \beta > 1$ is 'graded' in this quadrant. *Top Right*: Effect of increasing auto-repression strength, r, on the Fano-factor of the protein distribution in the auto-repression model. For the values chosen, both β and the Fano-factor go through maximum values at (different) intermediate values of r, before the distribution becomes sub-Poisson after the threshold value $r = r_0$; this happens exactly when $\beta = 0$. *Bottom Right*: Effect of increasing auto-repression strength r, on the protein distribution in the auto-repression model. Six different points from the above figure are chosen from the range where β remains positive. The other rates are the same as the ones used in the previous figure. r increases from lighter to darker values.

for $0 \le \lambda \le \lambda_{max}$, where a λ_{max} needs to be computed. If we can solve the above equations we have $\rho(\lambda)$ and the Poisson representation exists by construction.

Similarly, for the auto-repression model shown in Figure S1, if the Poisson representation exists, we can find the corresponding densities $\rho_\alpha(\lambda)$ that satisfy the Master Equations

$$\begin{aligned}\partial_t\rho_0(\lambda,t) &= -c_f\rho_0(\lambda,t) + c_b\rho_1(\lambda,t) + \partial_\lambda[\lambda\rho_0(\lambda,t)] \\ &\quad + r\,(\lambda\rho_1 - \partial_\lambda(\lambda\rho_1)) \qquad (5a)\\ \partial_t\rho_1(\lambda,t) &= c_f\rho_0(\lambda,t) - c_b\rho_1(\lambda,t) + \partial_\lambda[\lambda\rho_1(\lambda,t)] \\ &\quad - r\,(\lambda\rho_1 - \partial_\lambda(\lambda\rho_1)) - p_b\partial_\lambda\rho_1(\lambda,t)\,, \qquad (5b)\end{aligned}$$

with the boundary condition

$$-\,e^{-\lambda}\,(p_b - \lambda - r\lambda)\,\frac{\lambda^p}{p!}\,\rho_1(\lambda)\Big|_0^{\lambda_{max}} = 0 \tag{6}$$

for $0 \le \lambda \le \lambda_{max}$, where such a λ_{max} must be found.

III. RESULTS

To place our results for the auto-activation and auto-repression models in context, we will find it useful to compare these results with those derived for the linear pulsing model in [9]. Both auto-regulation models reduce to the linear pulsing model (LPM) in the limit where the auto-activation strength, a, or the auto-repression strength, r, tends to 0. In [9] we have shown how the 'phase-diagram' of all possible distributions for the linear pulsing model can be drawn in terms of the two scaled rates c_f/p_d and c_b/p_d.

A. Auto-activation

For the Auto-activation model (Figure S1) we have found a solution to the Master Equations (3), showing the validity of the superposition-of-Poissons representation, given below:

$$\rho(\lambda) = \mathcal{N}\, e^{\frac{a}{p_d+a}\lambda}\ \lambda^{\frac{c_f}{p_d+a}-1}\ \left(\frac{p_b}{p_d}-\lambda\right)^{\frac{c_b}{p_d+a}-1} \tag{7}$$

with $0 \le \lambda \le p_b/p_d$; $\mathcal{N}$ is the normalization constant. This exact expression leads naturally to the correct parametrization of the combinations of the rate constants that are relevant for analyzing this nonlinear model. We rescale λ by p_b/p_d so that it lies between 0 and 1. It is useful to rescale all rates by the effective protein degradation rate, p_d. For convenience in classifying the different kinds of protein distributions that arise in this model, we define the following parameters: $\alpha \equiv a\, p_b/(1+a)$, $\phi \equiv c_f/(1+a)$ and $\beta \equiv c_b/(1+a)$. We then have

$$\rho(\lambda) = \mathcal{N} e^{\alpha\,\lambda}\ \lambda^{\phi-1}\ (1-\lambda)^{\beta-1} \tag{8}$$

Note that ϕ and β characterize the singularity at the upper and lower limits of λ. Using this superposition-of-Poissons representation we have found that in each of the four 'quadrants' determined by ϕ and β greater or less than 1, the protein distribution has a distinct shape. Since the superposing density, $\rho(\lambda)$, is found to extend from $\lambda = 0$ to $\lambda = p_b$, $P(p)$ extends till $\sim p_b$. When the density diverges at both limits, i.e., ϕ and $\beta < 1$ yielding a ρ that is concave upwards the protein distribution is bimodal. When ρ vanishes at both limits, i.e., ϕ and $\beta > 1$, yielding a ρ that is concave downwards a broad bell-shaped distribution of proteins arises.

As the autoactivation strength $a \to 0$, $\alpha \to 0$, $\phi \to c_f$ and $\beta \to c_b$ the protein distribution of the autoactivation model becomes the exact steady-state distribution [9] obtained in the LPM. The latter is a Beta distribution

$$\rho(\lambda) = \mathcal{N}\ \lambda^{c_f-1}\ (1-\lambda)^{c_b-1} \tag{9}$$

Thus, the 'phase-diagram' of possible distributions in this model is very similar, in large regions of the parameter space ϕ and β, to that of the LPM, despite the auto-activation, once we identify ϕ and β in this model with c_f and c_b in the simple pulsing model.

We focus on the most interesting new feature that arises in this 'phase-diagram' in this model. Consider the quadrant where $\phi < 1$ and $\beta > 1$. When $a = 0$, i.e., in the LPM, we have found [9] long-tailed distributions with power-law behavior. In the auto-activation model, in contrast, two possibilities arise depending on whether α is lesser or greater than $\alpha_c \equiv (\sqrt{1-\phi}+\sqrt{\beta-1})^2$. In the former case long-tailed distributions with power-law regions arise, with an exponent $\phi-1$ as in the $a = 0$ case.

For $\alpha > \alpha_c$ the distribution becomes an unusually behaved *bimodal* distribution! To appreciate its nature we recall that when both ϕ and β are < 1 (Figure 1) bimodal distributions occur with the two modes always at 0 and p_b, i.e., at the edges of the allowed values of λ. As a the activation strength increases, the weights around 0 and p_b, are redistributed without affecting the separation between the modes. This is the classic 'binary response' [18] typically associated with auto-activation: cells may be thought to be divided into two sub-populations with low and high protein numbers and increasing activation strength only changes their relative proportions. In contrast, the new bimodal distribution exhibits a second mode not at p_b, the maximum allowed value of λ, but at intermediate values. As α is increased by increasing a, the protein distribution goes from being monotonically decreasing power-law to bimodal because auto-activation affects cells with intermediate numbers of proteins the most. Thus, when the feedback strength is strong enough that $\alpha > \alpha_c$, a new minimum and as well as a new maximum develop in $\rho(\lambda)$, at intermediate values of λ. Correspondingly, $P(p)$ becomes bimodal with the second mode arising at a value of $p < p_b$. As the activation strength increases, this mode tends to higher values of p, but the weight at 0 (the first mode) simultaneously erodes rapidly making the distribution effectively unimodal for strong enough activation. Thus in this quadrant even though bimodal distributions arise for intermediate activation strength, the response to increasing activation is really 'graded' as illustrated in Figure 1. As a increases, the protein distribution goes from being negatively skewed, with a large likelihood of obtaining a small number of proteins to a positively skewed distribution, with a large likelihood of obtaining a large number of proteins.

We point out the possible relevance of our results to the observation in a recent experiment of Maheshri et al. [26] of bimodal protein expression in a synthetic yeast system with positive feedback and no cooperativity as in our model. As the activation increases, their distribution goes from a broad bell-shaped distribution to the bimodal distribution similar to the one described above. Our model explains their observation of graded response of the 1xtetO promoter with increasing auto activation strength. As expected from our model, with increasing a, the mode at larger value travels further towards the right and acquires more weight until a Poisson like distribution

occurs.

Since this model is nonlinear the equations for all the moments are coupled and one needs the full distribution to obtain even the lowest two moments. Explicit expressions are displayed in the Supplementary section. These yield the Fano-factor (FF), the ratio of the variance to the mean of a distribution. FF may or may not go through a maximum value as a is increased, but beyond a threshold FF always decreases with a and tends to 1 as $a \to \infty$. Thus increasing auto-activation results in noise reduction, a role not conventionally associated with positive feedback. This is true since the gene is always "on" as the activation strength tends to infinity, and a Poisson protein distribution, with $FF = 1$ results. For any initial choice of parameters, for large enough a, both ϕ and β fall below 1, and the protein distribution becomes bimodal. However, in the limit $a \to \infty$, the mode at 0 is entirely eroded and $\rho(\lambda) \to \delta(p_b - \lambda)$: $P(p)$ becomes Poisson.

B. Auto-repression

The analysis of this model proceeds along the same lines as the auto-activation model. Once again, this formulation leads naturally to the correct parametrization of the combinations of the rate constants that are relevant for analyzing this nonlinear model. We define the new parameters, $\alpha \equiv r\,p_b/(1+r)^2$, $\phi \equiv c_f$ and $\beta \equiv p_b\,r/(1+r)^2 + (c_b - c_f\,r)/(1+r)$. All rates have been scaled by the protein degradation rate, p_d as before. In terms of these new variables, the steady state generating function is identical in form to that derived in the auto-activation model!

However, there is a subtle difference which has profound consequences: the parameter β can become *negative* for suitably chosen rates, p_b, c_f, c_b and r, in this model, unlike the auto-activation case. Thus the weighting probability density $\rho(\lambda)$ can be found, only if $\beta > 0$. This immediately implies that for $\beta > 0$, the protein distribution in the auto-repression model is super-Poisson, i.e., its FF is > 1 and thus 'noisier' than the Poisson distribution that arises in the simple birth-death model.

When β is < 0, we find that the protein distribution becomes sub-Poisson, i.e., its FFs becomes < 1. Thus, only when $\beta < 0$ can the auto-repression be said to be strong enough to cause reduction of the noise level in related models, such as the LPM and the auto-activation model. On analyzing the condition $\beta < 0$, we find that for any given value of the rates c_f, c_b and p_b, there is a threshold value of the repression strength, r_0, such that when r increases beyond this threshold value, the distribution becomes sub-Poisson (as illustrated in Figure 1). As seen in the top right panel of Figure 1 this suppression occurs for large values of r and over a narrow range. Exactly at the threshold value, the Fano-factor is found to be unity. The expression for r_0 is

$$r_0 = \frac{c_b + p_b - c_f + \sqrt{(c_b + p_b - c_f)^2 + 4\,c_f\,c_b}}{2\,c_f} \qquad (10)$$

For values of $r > r_0$, i.e., when the distribution is sub-Poisson, a formal expression for $\rho(\lambda)$ may be derived, with the understanding that it can no longer be interpreted as a probability density. In fact, λ now extends over the complex plane. Remarkably, even in this case, the functional form of $\rho(\lambda)$ remains the same for a suitably chosen contour. In this case, depending on whether c_f is < 1 or > 1, the protein distribution is a monotonically decreasing or a sharply peaked bell-shaped distribution, respectively.

When $\beta > 0$, we find that the different possible distributions of the auto-activation models all occur for auto-repression for appropriate values of α, β and ν, when $\beta > 0$, as illustrated in Figure 1. This underscores the inadvisability of naively inferring that the choice of auto-inhibition motif is designed to obtain noise suppression without further exploring the specific details of the system. See also [27] for a control and information theoretical perspective on the issue. Quantitatively, λ is in the range $0 \le \lambda \le p_b/(1+r)$ and so the protein distribution extends to about $p \sim p_b/(1+r)$. The effective parameter β is now a function of all the rates in the problem while the effective parameter $\phi = c_f$ as in the linear pulsing model.

IV. CONCLUDING REMARKS

The auto-regulation motif is ubiquitous in gene regulation [3, 17]. The auto-regulation models studied here are admittedly simplified descriptions of those observed in nature: we have not included separate transcription and translation steps. In prokaryotes, since mRNAs are rapidly translated into proteins, this is typically a reasonable approximation. For eukaryotic systems, when the mRNA time-scale is significant, these results should not be applied literally. Further, the effects of co-operative auto-regulation are not included in our models. However, even in this simple model a plethora of behaviors are observed including power laws, and bimodal distributions that behave in a graded fashion and sub-Poisson statistics. We have also established the utility of the Poisson Representation which yields quite naturally, the important, scaled, dimensionless parameters that characterize non-linear gene regulation models. We saw that auto-activation produces 'binary' responses to increasing activation strength and that auto-repression produces noise-suppressed sub-Poisson protein distributions in very limited regions of the parameter space. Our work serves to add a note of caution to assuming that positive and negative feedback, when found in natural biological systems, are present to serve these purposes.

V. ACKNOWLEDGEMENTS

CJ acknowledges support through contract HHSN272201000054C of NIAID. SIB wishes to acknowledge a useful discussion each with N. Maheshri, A. Walczak and N. Wingreen and several insightful discussions with R. R. Biswas.

[1] A. Eldar and M. Elowitz, Nature **467**, 167 (2010).
[2] J. Paulsson, Nature **427**, 415 (2004).
[3] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
[4] A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, and S. Tyagi, PLoS Biology **4**, 1707 (2006).
[5] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Cell **123**, 1025 (2005).
[6] J. M. Raser and E. K. O'Shea, Science **304**, 1811 (2004).
[7] T. B. Kepler and T. C. Elston, Biophys. J. **81**, 3116 (2001).
[8] J. E. M. Hornos, D. Schultz, G. C. P. Innocentini, J. Wang, A. M. Walczak, J. N. Onuchic, and P. G. Wolynes, Phys. Rev. E **72**, 059107 (2005).
[9] S. Iyer-Biswas, F. Hayot, and C. Jayaprakash, Phys. Rev. E **79**, 031911 (2009).
[10] J. M. Pedraza and J. Paulsson, Science **319**, 339 (2008).
[11] M. Savageau, Nature **252**, 546 (1974).
[12] A. Becskei and L. Serrano, Nature **405**, 590 (2000).
[13] M. Thattai and A. van Oudenaarden, Proc. Natl. Acad. Sci. USA **98**, 8614 (2001).
[14] Y. Dublanche, K. Michalodimitrakis, N. Kummerer, M. Foglierini, and L. Serrano, Mol. Syst. Biol. **2**, 41 (2006).
[15] D. Austin, M. Allen, J. McCollum, R. Dar, J. Wilgus, G. Sayler, N. Samatova, C. Cox, and M. Simpson, Nature **439**, 608 (2006).
[16] D. Thieffry, A. M. Huerta, E. Perez-Rueda, and J. Collado-Vides, Bioessays **20**, 433 (1998).
[17] U. Alon, An Introduction to Systems Biology: Design Principles of Biological Circuits (2006).
[18] A. Becskei, B. Seraphin, and L. Serrano, EMBO J. **20**, 2528 (2001).
[19] M. Acar, J. Mettetal, and A. van Oudenaarden, Nat. Genet. **40**, 471 (2008).
[20] W. Smits, C. Eschevins, K. Susanna, S. Bron, O. Kuipers, and L. W. Hamoen, Mol. Microbiol. **56**, 604 (2005).
[21] G. Suel, R. Kulkarni, J. Dworkin, J. Garcia-Ojalvo, and M. B. Elowitz, Science **315**, 1716 (2007).
[22] J. R. Klauder and E. C. G. Sudarshan (2006).
[23] J. Grandell (1997).
[24] D. Karlis and E. Xekalaki, International Statistical Review **73**, 35 (2005).
[25] C. W. Gardiner, Handbook of Stochastic Methods (1997).
[26] T.-L. To and N. Maheshri, Science **327**, 1142 (2010).
[27] I. Lestas, G. Vinnicombe, and J. Paulsson, Nature **467**, 163 (2010).

Supplementary Section for Mixed Poisson distributions in exact solutions of Stochastic Auto-regulation Models

1 Derivation of steady state distributions

1.1 Introduction

We introduce the method that we use to investigate the auto-activation and auto-repression models in the context of the linear pulsing model(LPM) where there is no feedback. This model is described by the reactions

$$\begin{aligned} D_{\text{off}} &\underset{c_b}{\overset{c_f}{\rightleftharpoons}} D_{\text{on}} \\ D_{\text{on}} &\xrightarrow{k_b} D_{\text{on}} + P \\ P &\xrightarrow{k_d} 0 \end{aligned} \tag{1}$$

Let the probabilities that the gene is in the on or off state at time t and exactly p proteins are present be denoted by $P_1(p,t)$ and $P_0(p,t)$, respectively. The Master Equations satisfied by P_0 and P_1 can be shown to be the following [1].

$$\begin{aligned} \frac{dP_0(p,t)}{dt} = & \quad -c_f P_0(p,t) \,+\, c_b P_1(p,t) \,+\, k_d\,[(p+1)P_0(p+1,t) - pP_0(p,t)\,] \\ \frac{dP_1(p,t)}{dt} = & \quad c_f P_0(p,t) \,-\, c_b P_1(p,t) \,+\, k_d\,[(p+1)P_1(p+1,t) - pP_1(p,t)\,] \\ & \qquad\qquad +k_b\,[\,P_1(p-1,t) \,-\, P_1(p,t)\,] \end{aligned}$$

One can solve this model using generating functions [1]. They are defined by $G(z,t) \equiv \sum_{p=0}^{\infty} P(p,t)\, z^p$ where $P(p,t) \,=\, P_0(p,t) + P_1(p,t)$. We solve the models with feedback using the Poisson representation and we will illustrate the method in the linear model. We assume the existence of probability densities $\rho_0(\lambda,t)$ and $\rho_1(\lambda,t)$ that yield $P_0(p,t)$ and $P_1(p,t)$ by

$$P_\alpha(p,t) \,=\, \int_0^\infty d\lambda\, \rho_\alpha(\lambda,t)\, e^{-\lambda}\, \frac{\lambda^p}{p!} \text{ where } \alpha = 0 \text{ or } 1. \tag{2}$$

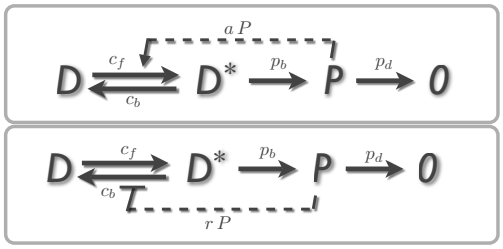

Figure 1: The auto-regulation models: The auto-activation/auto-repression model is shown on the top/bottom. D and D^* denote the gene in the 'off' and 'on' states, P denotes the protein. In both cases, auto-regulation is assumed to use monomers of the protein.

Note that the normalization condition $\sum_{p=0}^{\infty} P(p,t) = 1$ leads to the corresponding condition $\int d\lambda\, \rho(\lambda) = 1$ on $\rho(\lambda,t) \equiv \rho_0(\lambda,t) + \rho_1(\lambda,t)$. From the Master Equations for $P_0(p,t)$ and $P_1(p,t)$ we obtain the equations satisfied by the Poisson densities ρ_0 and ρ_1,

$$\partial_t \rho_0(\lambda,t) = -c_f \rho_0(\lambda,t) \,+\, c_b \rho_1(\lambda,t) \,+\, k_d\, \partial_\lambda[\lambda \rho_0(\lambda,t)] \tag{3}$$

$$\partial_t \rho_1(\lambda,t) = c_f \rho_0(\lambda,t) \,-\, c_b \rho_1(\lambda,t) \,+\, k_d\, \partial_\lambda[\lambda \rho_1(\lambda,t)] \,-\, k_b \partial_\lambda \rho_1(\lambda,t)\,. \tag{4}$$

In deducing these we impose the condition that the boundary terms resulting from integration by parts vanish:

$$k_d\, e^{-\lambda}\, (\lambda_{max} - \lambda)\, \frac{\lambda^m}{m!}\, \rho_1(\lambda) \Big|_0^{\lambda_{max}} = 0 \tag{5}$$

where λ lies in the interval $0 \le \lambda \le \lambda_{max}$. The boundary terms vanish identically at λ_{max}, provided that $\rho_1(\lambda)$ is less singular than $1/(\lambda_{max} - \lambda)$ and at the lower limit if ρ_1 vanishes at $\lambda = 0$. The solution we obtain $\rho_1(\lambda)$ does indeed behave as required and so the assumption that the boundary terms vanish can be justified a posteriori. We also have $\lambda_{max} = \frac{k_b}{k_d}$ as is physically reasonable since when the DNA is always on it leads to a Poisson distribution for the protein distribution with the λ-parameter equal to the maximum rate of protein production k_b/k_d. In the steady state, the coupled Master Equations for ρ_0 and ρ_1 can be solved and yield

$$\rho(\lambda) \,=\, \mathcal{N}\, \lambda^{-1+c_f}\, (k_b - \lambda)^{(-1+c_b)} \tag{6}$$

where the normalization factor $\mathcal{N}$ can be easily found by integrating over the allowed range of λ. Since $\rho_1 \propto \lambda\rho$ the vanishing of the boundary terms is easily justified. The result obtained is that in steady state the distribution $\rho(\lambda)$ is a scaled Beta Distribution with Beta parameters c_f and c_b and with λ ranging from 0 to k_b, i.e., $\lambda_{max} = k_b$. The steady state distribution $P(p)$, give by a superposition of Poisson distributions with the probability $\rho(\lambda)$ obtained above; using the integral representation of the confluent hypergeometric function

$$\int_0^u dx\, x^{\mu-1}\,(u-x)^{\nu-1}\,e^{\beta x} \;=\; \frac{\Gamma(\mu)\Gamma(\nu)}{\Gamma(\mu+\nu)}\,u^{\mu+\nu-1}\,{}_1F_1(\mu,\mu+\nu;\beta u) \tag{7}$$

we obtain

$$P(p) \;\propto\; {}_1F_1(c_f+p, c_f+c_b+p; -k_b)\,, \tag{8}$$

where the proportionality constant can once again be found from normalization. We follow this procedure to determine $P(p)$ without recourse to the standard generating function method for the models with positive and negative feedback below.

1.2 The Auto Activation Model

This model is described by the following reactions with the protein switching the gene from the off- to the on-state:

$$D \underset{c_b}{\overset{c_f}{\rightleftharpoons}} D^* \tag{9}$$

$$D + P \xrightarrow{a} D^* + P \tag{10}$$

$$D^* \xrightarrow{k_b} D^* \;+\; P \tag{11}$$

$$P \xrightarrow{k_d} \emptyset\,. \tag{12}$$

The protein probability distributions with the DNA in the D or D^* states denoted by $P_0(p,t)$ and $P_1(p,t)$ respectively satisfy the Master equations

$$\begin{aligned}\frac{dP_0(p,t)}{dt} &= -c_f P_0(p,t) \;+\; c_b P_1(p,t) \;-\; apP_0(p,t) \\ &\quad +\; k_d\,[(p+1)P_0(p+1,t) - pP_0(p,t)\,] \end{aligned}\tag{13}$$

$$\begin{aligned}\frac{dP_1(p,t)}{dt} &= c_f P_0(p,t) \;-\; c_b P_1(p,t) \;+\; apP_0(p,t) \;+\; k_d\,[(p+1)P_1(p+1,t) - pP_1(p,t)\,] \\ &\quad + k_b\,[\,P_1(p-1,t) \;-\; P_1(p,t)\,] \end{aligned}\tag{14}$$

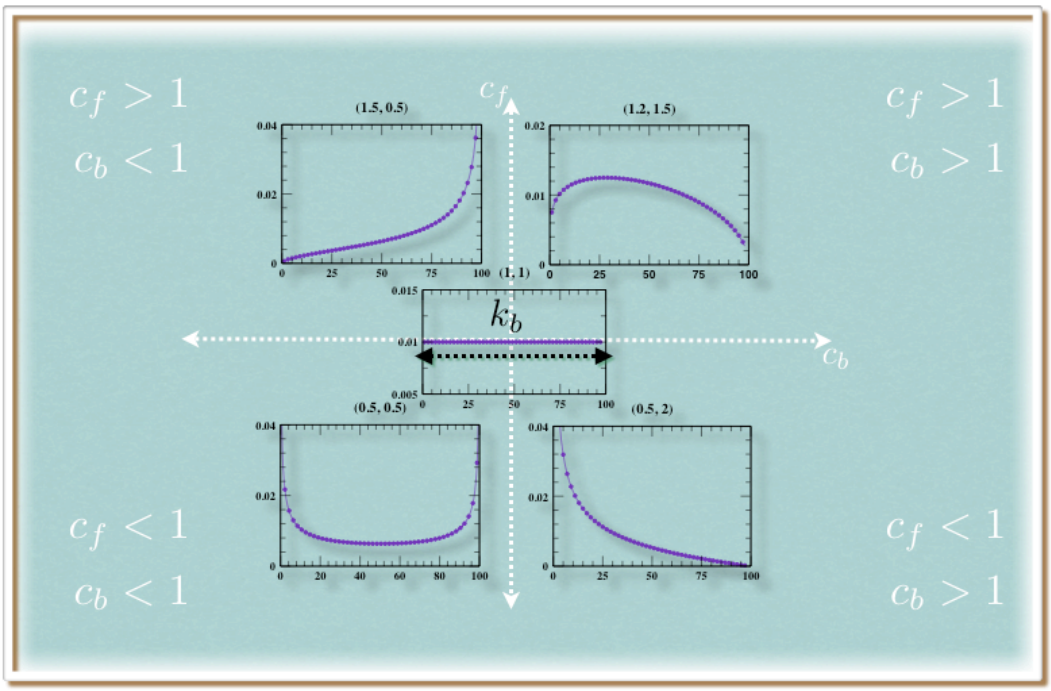

Figure 2: Steady state $\rho(\lambda)$ distributions for the LPM; they are labeled by (c_b, c_f) in units of k_d. The translation rate k_b is $100k_d$ for all the distributions. The figure shows prototype distributions for the four major regimes $c_f, c_b \lessgtr k_d$ in parameter space corresponding to the four quadrants. The case where all the regimes overlap with the uniform distribution for $c_f = c_b = k_d$ is also shown. All distributions extend from $\lambda = 0$ to $\lambda = k_b = 100$.

As before, we can rewrite the Master Equations for $P_\alpha(p,t)$ in terms of ρ_0 and ρ_1. The result is

$$\begin{aligned}\partial_t\rho_0(\lambda,t) &= -c_f\rho_0(\lambda,t) \,+\, c_b\rho_1(\lambda,t) \,+\, k_d\,\partial_\lambda[\lambda\rho_0(\lambda,t)] \\ &\quad - a\,(\lambda\rho_0 \,-\, \partial_\lambda(\lambda\rho_0)) \qquad (15)\\ \partial_t\rho_1(\lambda,t) &= c_f\rho_0(\lambda,t) \,-\, c_b\rho_1(\lambda,t) \,+\, k_d\,\partial_\lambda[\lambda\rho_1(\lambda,t)] \\ &\quad + a\,(\lambda\rho_0 \,-\, \partial_\lambda(\lambda\rho_0)) \,-\, k_b\partial_\lambda\rho_1(\lambda,t)\,. \qquad (16)\end{aligned}$$

with the requirement that the boundary terms vanish, i.e.,

$$\pm\, a\;\; e^{-\lambda}\,\frac{\lambda^m}{m!}\,\lambda\,\rho_0(\lambda,t)\Big|_0^{\lambda_{max}} \quad . \qquad (17)$$

where λ is restricted to the interval $0 \le \lambda \le \lambda_{max}$. Once again, the solution we find will satisfy these boundary conditions at $\lambda = 0$ and $\lambda_{max} = k_b/k_d$. Expressing all

rates in units of k_d we can solve the steady-state Master Equations to obtain

$$\rho(\lambda) \;=\; \mathcal{N}\, e^{\frac{a}{(1+a)}\lambda} \lambda^{\frac{c_f}{(1+a)}-1} \left(k_b - \lambda\right)^{\frac{c_b}{(1+a)}-1} \tag{18}$$

with $\lambda \in [0, k_b]$. We can use the probability density of $\rho(\lambda)$ to deduce the steady-state $P(p)$ directly without the use of generating functions, as before. We find

$$P(p) \;\propto\; {}_1F_1\left(\frac{c_f}{1+a} + p, \frac{c_f + c_b}{1+a} + p; -\frac{k_b}{1+a}\right), \tag{19}$$

where the proportionality constant is determined from normalization.

1.3 The Auto Repression Model

This model is described by the following reactions:

$$D \underset{c_b}{\overset{c_f}{\rightleftharpoons}} D^* \tag{20}$$

$$D^* + P \xrightarrow{r} D + P \tag{21}$$

$$D^* \xrightarrow{k_b} D^* \;+\; P \tag{22}$$

$$P \xrightarrow{k_d} \emptyset\,. \tag{23}$$

We can derive the Master Equations satisfied by the λ-densities as before.

$$\begin{aligned} \partial_t \rho_0(\lambda, t) = &-c_f \rho_0(\lambda, t) \;+\; c_b \rho_1(\lambda, t) \;+\; k_d\, \partial_\lambda[\lambda \rho_0(\lambda, t)] \\ &+\; r\; (\lambda \rho_1 \;-\; \partial_\lambda(\lambda \rho_1)) \end{aligned} \tag{24}$$

$$\begin{aligned} \partial_t \rho_1(\lambda, t) = &\; c_f \rho_0(\lambda, t) \;-\; c_b \rho_1(\lambda, t) \;+\; k_d\, \partial_\lambda[\lambda \rho_1(\lambda, t)] \\ &-\; r\; (\lambda \rho_1 \;-\; \partial_\lambda(\lambda \rho_1)) \;-\; k_b \partial_\lambda \rho_1(\lambda, t)\,. \end{aligned} \tag{25}$$

In this case, the requirement that boundary terms vanish translates to

$$e^{-\lambda} \left(k_b - k_d \lambda - r\lambda\right) \frac{\lambda^m}{m!}\, \rho_1(\lambda) \Big|_0^{\lambda_{max}}\,. \tag{26}$$

If we set $\lambda_{max} \;=\; k_b/(1+r)$ (using rate constants measured in units of k_d) the boundary terms can be made to vanish provided $\rho_1(\lambda)$ is well behaved at the limits. On solving these equations, we find that

$$\rho_1(\lambda) \;\propto\; e^{\frac{r}{1+r}\lambda}\; \lambda^{c_f}\; \left(k_b - \lambda(1+r)\right)^{\beta\,-1} \tag{27}$$

where we have defined

$$\beta \;\equiv\; \frac{c_b - r c_f + k_b r/(1+r)}{1+r}\,. \tag{28}$$

For $\beta > 0$, the solution is well behaved and the boundary conditions are satisfied; thus the description of the protein distribution as a superposition of Poisson representation is well-defined in this case. The density $\rho(\lambda)$ is found to be

$$\rho(\lambda) \;=\; \mathcal{N}\, e^{\frac{r}{1+r}\lambda}\; \lambda^{c_f-1}\; \left(\frac{k_b}{1+r} - \lambda\right)^{\beta-1} \quad \text{for } \beta > 0 \tag{29}$$

If $\beta \leq 0$ the boundary conditions are not satisfied and the representation as a superposition of Poisson distributions does not exist. Indeed, as will explicitly demonstrate later, $P(p)$ becomes sub-Poisson for $\beta < 0$ with a Fano factor less than 1. Therefore, it is not surprising that the protein distribution cannot be described by a superposition of Poisson distributions. However, formally we can write down a complex representation as used in quantum optics for squeezed states. [2] This follows from the fact that the integral representation of the confluent hypergeometric function used previously to obtain $P(m)$,

$$\int_0^u dx\, x^{\mu-1}\, (u-x)^{\nu-1}\, e^{\beta x} \;=\; B(\mu,\nu)\, u^{\mu+\nu-1}\, {}_1F_1(\nu, \mu+\nu; \beta\mu) \tag{30}$$

can be extended for $\mu < 0$ and $\nu < 0$ when the integral is taken along the Pochhammer contour. Using the Pochhammer contour, $\rho(\lambda)$ retains the same functional form even when $P(p)$ becomes sub-Poisson, i.e., when $\beta < 0$, but evidently it can no longer be interpreted as a probability density in this case. When convergence of boundary terms is imposed we can obtain $\rho(\lambda)$ from ρ_1 and thence derive the steady-state $P(p)$ directly as before. We find that, for all β

$$P(p) \;= \frac{\Gamma(c_f+p)\Gamma(c_f+\beta)}{\Gamma(c_f)\Gamma(c_f+\beta+p)} \left(\frac{k_b}{1+r}\right)^p \frac{1}{p!} \frac{{}_1F_1\left(c_f+m,\, \beta+c_f+p;\, -\frac{k_b}{(1+r)^2}\right)}{{}_1F_1\left(c_f,\, \beta+c_f;\, \frac{rk_b}{(1+r)^2}\right)}. \tag{31}$$

2 Exact expressions for the mean and the variance

Since both the auto-activation and auto-repression models are non-linear feedback models, it is not possible to derive the exact mean of the protein distribution from the rate equations, since such a description ignores the correlation between the promoter and protein dynamics. As is well-known because of the non-linearity, the Master Equation cannot be used to derive an evolution equation for just the mean (in contrast to linear models): the equation for the mean is coupled to higher moments leading to a hierarchy of coupled equations that cannot be truncated if one desires an exact expression even for the first moment; the entire probability distribution needs to be explicitly evaluated to obtain even the mean. From the exact distribution that we have derived we can obtain the mean and the variance.

2.1 The Auto activation Model

The steady-state mean, μ, is seen to be the first moment of $\rho(\lambda)$ and is given below:

$$\mu \equiv \langle p \rangle = \frac{{}_1F_1\left(\frac{c_f}{a+1}+1; \frac{a+c_b+c_f+1}{a+1}; \frac{ak_b}{a+1}\right)}{{}_1F_1\left(\frac{c_f}{a+1}; \frac{c_b+c_f}{a+1}; \frac{ak_b}{a+1}\right)} \frac{c_f k_b}{(c_b+c_f)} \tag{32}$$

We can also obtain from the second moment of $\rho(\lambda)$

$$\langle p(p-1) \rangle = \frac{{}_1F_1\left(\frac{c_f}{a+1}+2; \frac{c_b+c_f}{a+1}+2; \frac{ak_b}{a+1}\right) c_f\,(a+c_f+1)\,k_b^2}{{}_1F_1\left(\frac{c_f}{a+1}; \frac{c_b+c_f}{a+1}; \frac{ak_b}{a+1}\right)(c_b+c_f)\,(a+c_b+c_f+1)}\,. \tag{33}$$

The variance $\sigma^2 = \langle p^2 \rangle - \mu^2$ can be written down from the above expressions. As before, we have expressed all rates in units of the decay rate k_d. In the limit $a \to 0$, the model reduces to the LPM and indeed these expressions reduce to the simple results for the pulsing model[1],

$$\begin{aligned} \mu_0 &= \frac{c_f k_b}{c_b + c_f} \text{ and} \\ \sigma_0^2 &= \mu_0 \left(1 + \frac{c_b k_b}{(c_b+c_f)(c_b+c_f+1)}\right) \end{aligned} \tag{34}$$

The first term the variance is the noise in the simple Poisson process leading to a Fano factor of unity while the second term in the expression above corresponds to the additional noise from the DNA flipping between the on and off states.

2.2 The Auto Repression Model

In this model, the expression for the mean is

$$\mu = \frac{k_b c_f\, {}_1\tilde{F}_1\left(c_f+1; \frac{c_b+c_f+r\left(c_b+c_f+k_b\right)}{(r+1)^2}+1; \frac{rk_b}{(r+1)^2}\right)}{(r+1)\, {}_1\tilde{F}_1\left(c_f; \frac{c_b+c_f+r\left(c_b+c_f+k_b\right)}{(r+1)^2}; \frac{rk_b}{(r+1)^2}\right)} \tag{35}$$

where ${}_1\tilde{F}_1$ is the regularized confluent hypergeometric function. The expression for the second moment is

$$\nu = \frac{k_b c_f \left(k_b + r + 1\right) {}_1\tilde{F}_1\left(c_f + 1; \frac{c_b + c_f + r\left(c_b + c_f + k_b\right)}{(r+1)^2} + 1; \frac{r k_b}{(r+1)^2}\right)}{(r+1)^2 {}_1\tilde{F}_1\left(c_f; \frac{c_b + c_f + r\left(c_b + c_f + k_b\right)}{(r+1)^2}; \frac{r k_b}{(r+1)^2}\right)} - \frac{k_b^2 c_f \left(r\left(c_b + k_b - c_f\right) + c_b + r^2\left(-c_f\right)\right) {}_1\tilde{F}_1\left(c_f + 1; \frac{c_b + c_f + r\left(c_b + c_f + k_b\right)}{(r+1)^2} + 2; \frac{r k_b}{(r+1)^2}\right)}{(r+1)^4 {}_1\tilde{F}_1\left(c_f; \frac{c_b + c_f + r\left(c_b + c_f + k_b\right)}{(r+1)^2}; \frac{r k_b}{(r+1)^2}\right)} \tag{36}$$

and from these expressions, the variance σ^2 can be written down using $\sigma^2 = \nu - \mu^2$. Once again, the LPM results for the mean and variance are recovered in the limit when the feedback strength becomes 0, i.e. as $r \to 0$.

3 The phase diagram of allowed steady state distributions

From the functional form of $\rho(\lambda)$, the important dimensionless combinations (ratios) of the rates in the model can be deduced and they can be used to demarcate the regions of the parameter space of the model where qualitatively distinct distributions occur.

3.1 The linear pulsing model

Once again, it will be useful to compare the results for the auto-activation and auto-repression models with that obtained for the linear pulsing model and so we recall them here. Recall that for this case $\rho(\lambda) = \mathcal{N}\,\lambda^{-1+c_f}\,(k_b - \lambda)^{(-1+c_b)}$. First, the shape of the distribution is evidently determined by c_f and c_b; k_b is just a scaling variable that determines the extent of the distribution. By examining the possible extrema of $\rho(\lambda)$ for the shape determining variables c_f and c_b, it can be seen that the parameter space can be divided into the four 'quadrants' shown in Figure(2). In the first quadrant (top, right), $\rho(\lambda)$ has one maximum at $0 < \lambda < k_b$ and thus the resultant $P(p)$ is a broad bell-shaped curve. In the second quadrant (top, left), $\rho(\lambda)$ diverges at $\lambda = k_b$ and so the corresponding $P(p)$ is a narrow Poisson like distribution with a bell-shaped background. Indeed, as $c_b \to 0$, the distribution becomes Poisson. In the third quadrant (bottom, left), $\rho(\lambda)$ is U shaped diverging at both ends with a minimum at $0 < \lambda < k_b$. As a result, $P(p)$ is bimodal with modes around 0 and k_b. Finally, in the fourth quadrant, i.e., the bottom right, $\rho(\lambda)$ is monotonically decreasing with a

maximum at 0 and a power law tail with an exponent $c_f/k_d - 1$ and the resultant $P(p)$ is also monotonically decreasing with a power law tail with an exponent $c_f/k_d - 1$.

3.2 The Auto activation Model

For this case, it was shown that the Poisson density is given by $\rho(\lambda) = \mathcal{N} e^{\frac{a}{(1+a)\lambda}} \lambda^{\frac{c_f}{(1+a)}-1} (k_b - \lambda)^{\frac{c_b}{(1+a)}-1}$ and $0 \leq \lambda \leq k_b$. The extent of the distribution is thus determined by k_b. By examining the extrema of the function, it can be shown that the shapes of the distributions in this model are determined by $\phi \equiv c_f/(1+a)$ and $\beta \equiv c_b/(1+a)$ and $\alpha \equiv ak_b/(1+a)$. Qualitatively, the parameter space can still be divided up into four quadrants in terms of (ϕ, β) similar to it was in terms of (c_f, c_b) in the LPM case. This is shown in Figure(3). Qualitatively, the behaviors of $\rho(\lambda)$ and $P(p)$ in the first, second and third quadrants remain the same as that for the LPM case, if the parameter space is characterized by (ϕ, β) in the place of (c_f, c_b) in the linear pulsing model. In the fourth quadrant, the behavior of $\rho(\lambda)$ becomes α dependent as shown in Figure(4). $\rho(\lambda)$ is monotonically decreasing as before with a power law exponent equal to $\phi/k_d - 1$ for $\alpha < (\sqrt{1-\phi} + \sqrt{\beta - 1}\,)^2$. For $\alpha = (\sqrt{1-\phi} + \sqrt{\beta - 1}\,)^2$, $\rho(\lambda)$ develops a point of inflection at intermediate λ. For $\alpha > (\sqrt{1-\phi} + \sqrt{\beta - 1}\,)^2$, $\rho(\lambda)$ has two modes, one at 0 and the other at a value of $\lambda < k_b$. In this case, the system develops a new mode at a finite value of λ which is less than the edge of the interval at k_b, in contrast to the the bimodal in the third quadrant. As the value of α is increased, the second mode tends toward k_b. When $\alpha \to \infty$, the weight at the mode at 0 tends to 0 and the second mode tends to k_b and the distribution tends to a Poisson with mean k_b/k_d. We emphasize that this new kind of bimodal is not present in the LPM and is a result of the feedback included in the model. [1]

References

[1] S. Iyer-Biswas, F. Hayot and C. Jayaprakash, Phys. Rev. E (2009), 79, 031911

[2] John R. Klauder and E. C. G. Sudarshan, "Fundamentals of Quantum Optics", Dover Publications (2006),

[1] We note that in the third quadrant, when c_f, c_b are themselves $< k_d$, true bimodal distributions are obtained for $P(p)$ but if $c_f, c_b > k_d$ such that ϕ, β are still in the third quadrant, then $\rho(\lambda)$ has a minimum very close to 0 and so $P(p)$ has a small delta function mode at 0 and is nominally a bimodal, but for all practical purposes has a mode just at k_b.

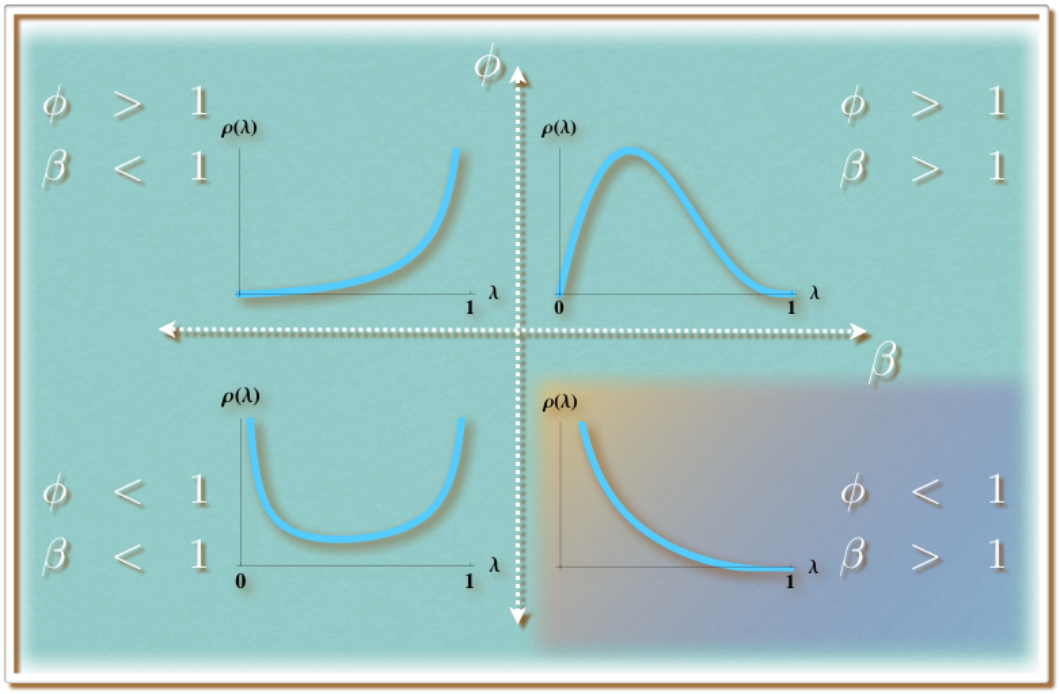

Figure 3: The phase diagram of the superposing density $\rho(\lambda)$ for the Auto-activation model, in terms of the variables ϕ and β defined in the text.
[Notice how strongly reminiscent this phase diagram is of the corresponding figure for the LPM, in terms of the variables c_f and c_b. Since this figure is meant to be illustrative, the actual values of the parameters used to generate the distribution in each quadrant are not important. The variable λ has been rescaled by p_b/p_d so that the distribution extends from $\lambda = 0$ to $\lambda = 1$.]

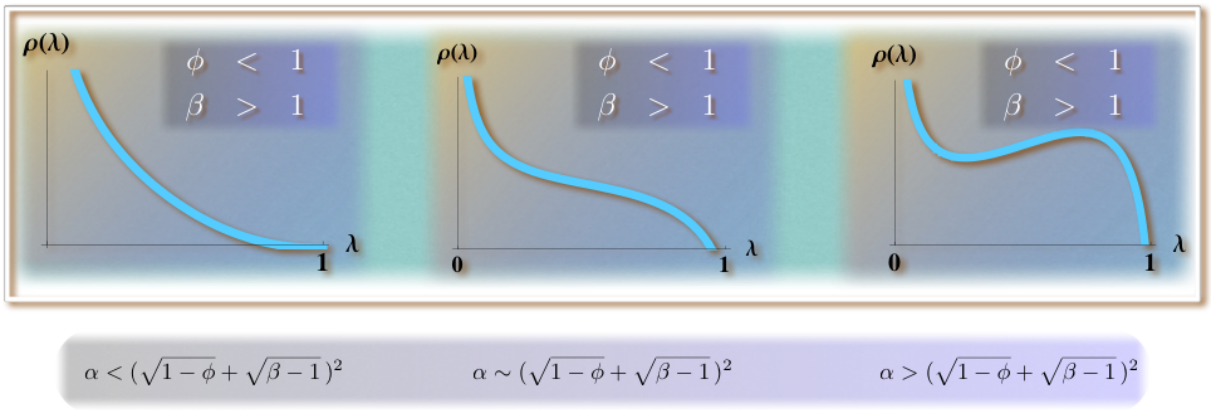

Figure 4: Appearance of a new kind of bimodal distribution in the Auto-activation model.
[While Figure 3 shows how strongly reminiscent distributions obtainable in the Auto-activation model are of those in the LPM, a closer look at the 'fourth-quadrant' in Figure 3 reveals that a new kind of bimodal distribution appears when appropriate conditions are satisfied by the parameters ($\alpha > (\sqrt{1-\phi} + \sqrt{\beta - 1}\,)^2$), while still in this quadrant. The leftmost distribution is the same as the one shown in the 'fourth-quadrant' of Figure 3. As a is increased (from left to right), such that the $\phi < 1$ and $\beta > 1$, a new mode develops at intermediate values of λ (< 1) when the above inequality is satisfied.]